AD-A265 730

CUMENTATION PAGE

Form Approved
OMB No. 0704-0188

2

2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
May 1993	Final Report - 11/1/90 - 10/31/92

4. TITLE AND SUBTITLE
Development of Practical MO Techniques for Prediction of the Properties and Behavior of Materials

5. FUNDING NUMBERS
61102F 2303 B2

6. AUTHOR(S)
Michael J. S. Dewar

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
University of Florida
219 Grinter Hall
Gainesville, FL 32611

8. PERFORMING ORGANIZATION REPORT NUMBER
AFOSR-TR- 93 0410

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
AFOSR/NC
Building 410, Bolling AFB DC
20332-6448

10. SPONSORING/MONITORING AGENCY REPORT NUMBER
AFOSR-91-0085

DTIC
ELECTE
JUN 14 1993
S B D

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION/AVAILABILITY STATEMENT
APPROVED FOR PUBLIC RELEASE; DISTRIBUTION IS UNLIMITED.

12b. DISTRIBUTION CODE

13. ABSTRACT (Maximum 200 words)

The original SAM1 program has been completely rewritten and optimized, and geometry optimization is now carried out using analytical derivatives instead of finite difference. Satisfactory parameters for carbon and hydrogen were obtained. However, problems were found for the case of nitrogen and oxygen, particularly for compounds containing N-N bonds. Testing of SAM1 was carried out by performing calculations for an extensive set of molecular species for which apparently reliable experimental data are available.

93-13290

14. SUBJECT TERMS

15. NUMBER OF PAGES
5

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
UNCLASSIFIED	UNCLASSIFIED	UNCLASSIFIED	

GRANT AFOSR-91-0085. FINAL REPORT.

Michael J. S. Dewar, Principal Investigator

INTRODUCTION

A major concern of my research group over the last twenty years has been the development of quantitative quantum mechanical procedures that chemists, including ourselves, could use as a practical aid in studies of chemical problems, in the same kind of way as NMR or mass spectrometry. To be useful in this connection, a procedure must reproduce the energies and other relevant properties of molecules with "chemical" accuracy, and it must also be applicable to the molecules in which chemists are specifically interested, which are often quite large, at reasonable cost, using readily available computers.

It seemed clear to me from the start that current so-called ab initio methods would not be able to meet these conditions in the foreseeable future and this indeed is still the case today [1]. We therefore adopted an expedient which has proved succesful in many other areas of theoretical chemistry where exact solutions of key equations are unavailable or too expensive, i.e. taking a crude (and hence cheap) approximation and then trying to upgrade its accuracy by introducing adjustable parameters. The Debye-Hückel theory of strong electrolytes is a classic example.

This so-called *semiempirical* approach to quantum chemistry was not in itself novel. Indeed, the approximations we have used were first introduced by Pople. However, previous attempts to parametrize them had failed to give satisfactory results and it had become generally accepted that procedures of this kind could never be of any real value in chemistry. We were able to show that this failure was due simply to lack of effort [2]. Over the years we have been able to develop a series of successively better treatments, based on Pople's INDO and MNDO approximations, and the three latest, MINDO/3 [3], MNDO [4], and AM1 [5], are now being widely used. Our own studies of a wide range of problems concerning chemical behaviour, including the mechanisms of numerous reactions, have led in many cases to major revisions of conclusions that had become embedded in chemical theory [1,6,7].

The success of these treatments is admittedly astonishing, given the extreme crudity of the approximations on which they are based. Many theoreticians have indeed rejected our claims as mathematically impossible, an attitude which be wholly justifiable if our treatments had been put forward as approximations to the Schrödinger equation. This, however, is not the case. As we have repeatedly pointed out, our object has been quite different. We have been trying to develop effective molecular models [6a,6d]. A model is a simple device that mimics the behaviour of a more complex one. We may then be able to predict the behaviour of a device that is too complex for rigorous analysis by observing the behaviour of an effective model. The value of a model in this connection depends only on how well it mimics the behaviour of the parent system, not on direct comparisons of the two. One cannot, for example, dismiss a model on the grounds that it is "made of cheap plastic". Assessment of our treatments must likewise be based solely on how well they perform in practice, not on the accuracy of the approximations on which they are based.

In any MO treatment of a molecule, each of the terms in the expression for the total energy can be given a physical interpretation, corresponding to the kinetic energy of one of the particles 1(electrons or nuclei) involved or to the electrostatic interaction between two of them. Errors in

the energies calculated for a number of different molecules can then usually be attributed to errors in specific terms. In our approach, where some or all of these terms are replaced by parametric expressions, the errors can then be corrected, or at least reduced, by appropriate adjustment of the corresponding parameters.

The extent to which the inherent deficiencies of a given MO treatment can be countered by parametrization is naturally limited. One cannot for example hope to eliminate errors that are directly due to the basic approximations made in the treatment, for example the neglect of overlap in treatments such as INDO or NDDO. Semiempirical treatments should therefore be based on the best possible approximations. Here, however, limits are set by the performance of currently asavailable computers. Since our object is to predict the behaviour of molecules of interest to chemists,which are commonly quite large, we must be able to carry out calculations for the corresponding models.

Parametrization involves finding a minimum in the *parameter hypersurface*, the multidimensional hypersurface that represents the RMS error in the properties calculated for the basis set molecules as a function of the parameters. The parameters are then tested by carrying out calculations for a selected set of molecules for which experimental data are available, including as many different kinds of molecule as possible. This test involves a subjective assessment of the significance of the errors for specific molecules. If the results for one or more key molecules are unsatisfactory, the parametrization is repeated with increased weight being given to molecules of that kind. The situation is further complicated by the fact that parameter hypersurfaces commonly have numerous minima. There is no way to tell whether a given minimum is, or is not, the global minimum. Furthermore, there is no guarantee that the global minimum for a given basis set will correspond to the "best" set of parameters for the atoms in question because it may lead to chemically unacceptable errors in the results for specific molecules. In short, parametrizing one of our semiempirical procedures is an extremely laborious and very frustrating undertaking which moreover requires chemical knowledge and judgement as well as perseverence.

Our two latest treatments, MNDO and AM1, are based on Pople's NDDO approximation, further simplified to save computing time. These additions include the core approximation, estimation of the electron repulsion integrals from a simple parametric function suggested by Dewar and Sabelli, and allowance for electron correlation by the Pariser-Parr [8] expedient of adjusting the electron repulsion integrals (EE).

MNDO suffered from a systematic overestimation of the repulsions between nonbonded atoms which led to an overestimation of steric effects and intermolecular repulsions, leading in particular to failure to reproduce hydrogen bonds. When attempts to correct these errors by reparametrization failed, we tried modifying the core repulsion function by adding terms containing additional parameters. This expedient led to an improved treatment (AM1) in which the systematic errors in MNDO seemed at first to have been overcome. Extensive use of AM1 has, however, shown that although the situation in AM1 is better, the errors in question have not been eliminated. In particular, although AM1 reproduces the heats of formation of hydrogen bonds reasonably well, the predicted geometries are unsatisfactory.

We have now discovered the real cause of the error. The parametric function used to calculate the electron-electron repulsions (EE) in MNDO and AM1 does not have the proper dependence on internuclear distance, the calculated repulsions between electrons on two different atoms being too small at distances significantly greater than the normal bond length. Since the core-

electron attractions are equated to a sum of EE-type terms, they too are underestimated, and since the interactions between two atoms involve *one* set of electron-electron repulsions but *two* sets of core-electron attractions, the error in the EE integrals leads to a underestimation of the attractions between the atoms and hence to a total energy that is too positive. The error in MNDO was therefore due to underestimation of the electron-electron repulsions, which, in the MNDO formalism, are unrelated to the core repulsions. Any improvement brought about by reparametrization, or by changes in the CR function, should therefore be achieved only at the expense of compensating errors elsewhere, the latter being greater, the more effectively the "MNDO errors" are dealt with. Since the parametrization of AM1 was designed to optimize its overall performance, the MNDO errors were reduced but not completely eliminated.

This conclusion is supported by Stewart's PM3 procedure [9], a reparametrization of AM1 directed to eliminating these errors completely. It now seems to be generally agreed that any gains achieved in this way are counterbalanced by more serious losses elsewhere.

This problem might be overcome by modifying the empirical EE function in AM1. However, any such modification might well lead to other problems. A better solution would be to attack the problem at its root by using theoretical values for the EE integrals, in other words, by replacing the basic AM1 approximation by a better one in which the simplification leading to a specific type of error is avoided. This approach would have the further advantage of allowing easy extension to spd basis sets, the AM1 formalism becoming very cumbersome if d AOs are included.

Work on this new treatment, which we termed SAM1 (Semi-Ab-initio Model 1) began some years ago at the University of Texas, in Austin. A corresponding computer program was developed and included in a version AMPAC and a stsrt was made on parametrizing it for the "organic" elements (C,H,O,N). The work carried out with support from the above AFOSR grant has been an extension of this work.

CURRENT STATUS OF SAM1

SAM1 follows the same basic pattern as AM1, being likewise based on the NDDO approximation together with the core approximation, and the one-center integrals are likewise treated as paraameters. The electron repulsion integrals are calculated theoretically, using the STO-3G basis set, and scaled by a measure of orbital overlap to allow for electron correlation, following the idea pioneered by Pariser and Parr [8]. The field due to the core of atom m is equated to *minus* that generated by Z_m valence shell s electrons, Z_m being the core charge in units of the electronic charge. Thus the attraction between the core of atom m and an electron in the AO ϕ_n of atom n is set equal to $-Z_m(s_m s_m;\phi_n\phi_n)$ where $(s_m s_m;\phi_n\phi_n)$ is the electron repulsion integral between the s AO of atom m and the AO ϕ_n of atom n.

Work on SAM1 began three years ago in Austin, at the University of Texas, before I moved to Florida. We wrote a basic computer program to carry out SAM1 calculations, which was incorporated in the current version of our AMPAC and parametrization programs, and preliminary parametrizations established what seemed to be a suitable form for the Pariser-Parr scaling function. The move to the University of Florida naturally caused disruption, particulary since it involved transferring our programs from the Alliant FX8 computer we had in Austin to the two SUN work stations, and an IBM RSC 6000 workstation, which we acquired in Gainesville. It is interesting to note that these together have provided us with at least ten times more computing

time than the Alliant while costing, collectively, less than one-fifth as much. A brief summary of the status of SAM1 at the end of 1991follows.

A. Computer Programs. The original SAM1 program was a makeshift affair in which the integrals were taken from a standard package and in which geometry optimization was carried out using derivatives found by finite difference. It has now been completely rewritten and optimized and geometry optimization is now carried out using analytical derivatives.

B. Parametrization. Preliminary studies confirmed the efficacy of the SAM1 algorithm established in Austin, in particular the scaling function used to allow for electron correlation. Parameters for carbon and hydrogen were obtained which gave results as good as those from AM1, using the unmodified MNDO CR function. While we later included Gaussian terms to improve the calculated geometries, the number of parameters for C and H is still less in SAM1 than in AM1.

Problems arose in the case of nitrogen and oxygen, particularly for compounds containing NN bonds for which the calculated heats of formation were too negative. Repeated attempts to improve the results led to increased errors elsewhere. We are now sure that the error is due to to another basic simplification in the SAM1 formalism, namely use of the ZDO approximation. As is well known, this leads to an underestimation of the exchange repulsions between filled orbitals and the errors are particularly large for filled AOs, i.e. for lone pairs. We have therefore abandoned attempts to improve the situation by further modifcation of the parameters.

C. Results. SAM1 has been tested by carrying out calculations for an extensive set of molecular species for which apparently reliable experimental data are available. Tables 1 compares the results for organic (CHON) molecules, given by SAM1, AM1, and PM3. The quantities listed are mean unsigned (MU) errors, root mean square (RMS) errors, and mean signed (MS) in the calculated heats of formation.

TABLE 1. NEAN ERRORS (KCAL/MOL) IN CHEATS OF FORMATION

Number of Molecules	Type of Error	Procedure		
		SAM1	AM1	PM3
217	MU	4.16	7.34	4.44
	RMS	5.66	10.30	6.07
	MS	-0.17	0.56	-0.27

Better results are expected when parametrization of SAM1 is finalized.

Work on this project is being continued with further support from AFOSR.

REFERENCES

1. M. J. S. Dewar, *Int. J. Quantum Chem.* 1992, **44**, 427.

2. M. J. S. Dewar, *J. Mol. Struct.* 1983, **100**, 41.

3. R. C. Bingham, M. J. S. Dewar, and D. H. Lo, *J. Am. Chem. Soc.* 1975, **97**, 1285, 1294, 1302, 1307.

4. M. J. S. Dewar, and W. Thiel, *J. Am. Chem. Soc.* 1977, **99**, 4899, 4907.

5. M. J. S. Dewar, E. G. Zoebisch, E. F. Healy, and J. J. P. Stewart, *J. Am. Chem. Soc.* 1985, **107**, 3902.

6. For a general account and references, see Dewar, M. J. S. *A Semiempirical Life;* American Chemical Society, Washington, DC, 1992. The following papers provide summaries of specific appl;ications: M. J. S. Dewar, (a) *J. Am. Chem. Soc.* 1983, **105**, 209; (b) *ibid.* 1984, **106**, 669; (c) *Enzyme,* 1986, **36**, 8; (d) *Int. J. Quantum Chem.; Quantum Chem. Symp.* 1988. **22**, 557; (e) *J. Mol. Struct. Theochem.* 1989, **200**, 301.

7. See e.g. M. J. S. Dewar and M. L. McKee, *Pure & Applied Chem.* 1980, **52**, 1431; M. J. S. Dewar and K. M. Merz, (a) *J. Am. Chem. Soc.* 1984, **106**, 4040; (b) *ibid.* 1986, **108**, 5142; (c) *ibid.* 1986, **108**, 5146; (c) *ibid.* 1986, **108**, 5634; (d) *ibid.* 1986, **108**, 6553; M. J. S. Dewar and J. M. Ruiz, *Tetrahedron,* 1987, **43**, 2661; M. J. S. Dewar and D. M. Storch. *J. Chem. Soc. Perkin Trans. II,* 1989, 877; M. J. S. Dewar, J. C. Hwang, and D. R. Kuhn, *J. Am. Chem. Soc.* 1991, **113**, 2088; M. J. S. Dewar and K. M. Dieter, *Biochemistry,* 1988, **27**, 3302; M. J. S. Dewar and M. L. Krull, *J. Chem. Soc. Chem. Commun.* **1990,** 333; M. J. S. Dewar and C. Jie, *(a) J. Am. Chem. Soc.* 1986, **108**, 5893; (b) *ibid.* 1989, **111**, 511; (c) *Tetrahedron* 1988. **44**, 1351; (c) *Acc. Chem. Res.* 1992, **11**, 537.

8. R. Pariser and R. G. Parr, *J. Chem. Phys.* 1953, **21**, 466, 767.

Michael J. S. Dewar

Michael J. S. Dewar, Principal Investigator